Knowledge in the Head and on the Web: Using Topic Expertise to Aid Search

Geoffrey B. Duggan
Manchester Business School,
University of Manchester,
Manchester, M15 6PB
geoffrey.duggan@manchester.ac.uk

Stephen J. Payne
Manchester Business School,
University of Manchester,
Manchester, M15 6PB
stephen.payne@manchester.ac.uk

ABSTRACT
The importance of background knowledge for effective searching on the Web is not well understood. Participants were given trivia questions on two topics and asked to answer them first using background knowledge and second by searching on the Web. Knowledge of a topic predicted search performance on that topic for all questions and, more importantly, for questions for which participants did not already know the answer. In terms of process, greater topic knowledge led to less time being spent on each Webpage, faster decisions to give up a line of inquiry and shorter queries being entered into the search engine. A more complete theory-led understanding of these effects would assist workers in a whole range of Web-related professions.

Author Keywords
Web search, query formulation, information scent, navigation, domain knowledge, expertise.

ACM Classification Keywords
H5.4 Information interfaces and presentation (e.g., HCI): Hypertext/Hypermedia.

INTRODUCTION
Is the role of acquired declarative knowledge in human expertise radically undermined by the extraordinary availability of information on the Web and the success of search engines? Or does knowledge in the head allow more successful use of these external knowledge tools? The second of these two questions can be addressed empirically. In this article we introduce a simple experimental design for addressing this question and for shedding light on the processes by which background knowledge may affect Web search.

The process of searching through an electronic database has been extensively examined within the Library and Information Sciences literature [12, 15] and in recent years the field has begun to focus on the specific challenges posed by the Web. One method of inquiry has been the analysis of log transactions from search engines [9, 10]. This provides a very large dataset for the researcher to analyse and enables a representative profile of typical query formulation and initial page selection to be developed. One limitation of many studies is that they ignore navigation subsequent to the search query. This was emphasised by Mat-Hassan and Levene [16], who contrasted the two processes in search of an online documentation system and found that users were more likely to rely on new or modified queries than navigation by link-following.

More generally, studies of very large, naturally occurring datasets struggle to capture the individual context associated with each particular search. Variables like the users' goal, search experience and knowledge of the topic will all affect searching behaviour yet are difficult to extract from log transactions. Laboratory based studies have provided a means to understand the influence of such factors. For example, Hsieh-Yee [7] found that increased search expertise (librarians vs. novice searchers) led to more synonymous terms, greater use of the thesaurus and more combinations of search terms when searching within a library database. Similar effects on search tactics were produced with increased subject or domain knowledge but only for the librarians and not within the group of novice searchers.

Search expertise and domain knowledge were also manipulated by Hölscher and Strube [6] and their tasks *were* completed using the World Wide Web. Search expertise was determined by interview and pre-test whereas the domain experts were undergraduates in the topic of the search task (Economics). Like [7] they found that both variables affected search behaviour and a combination of both types of expertise led to improvements in query selection and general query formatting. There also appeared to be an overall benefit of both variables on success in finding the target information although it is unclear whether this effect was reliable. Finally, it is worth noting the study was conducted at least 8 years ago and the failure rate while searching was high, prompting the authors to conclude there was "much room for improvement in Web-based searching" (p. 345), and prompting us to conclude that the basic phenomena they report are worthy of further controlled study.

Both [7] and [6] have been influential and have more recently been followed by other studies looking at the *process* of search and the role of domain expertise. Higher domain knowledge was found to produce more effective search query formulation in a simulated database environment [5] and within an actual microbiology database [20]. Similarly, general experience searching the World Wide Web predicts faster and more successful location of target Web sites [13].

Another important study using the World Wide Web [1] conflated search and domain expertise by comparing healthcare and online shopping experts. Both types of expert were not only knowledgeable in their specialist topic but, crucially, also very experienced at using the Web to retrieve information about that topic. When tested, all nine participants were more proficient within their area of expertise than the other topic. This was chiefly achieved by accessing more specialist Websites and exploiting their domain knowledge once there. For example, experienced shoppers knew which Websites had the lowest prices and whether this information would need to be verified or not.

One observation that arises from this study is the need to acknowledge that the Web is distinct from other electronic databases, not least because of its size and the unstructured way in which it has been constructed, but also because of the unreliability of some of its information. This means that some of the expertise associated with finding information consists of simply knowing the right Web sites to visit. In [1] experts just typed in the URL of the most useful Websites. Although this behaviour is interesting and representative of much Web usage it is not our concern in this article. Rather, we are interested in the way that users search for information when they do not necessarily know the most appropriate Websites to check. Furthermore, given that domain expertise has consistently been shown to affect searching within other electronic databases [5,7,20] and on the Web 8 years ago [6] we wish to test both whether and how domain expertise can assist searching on the Web today.

To this end we shall adapt and develop a methodology introduced by Wildemuth [20]. Her study investigated whether microbiology students' search tactics changed as their topic knowledge increased over an academic year. To address this, students were asked questions about the topic at 3 separate junctures several months apart. The mean knowledge score across the entire group from a test session was then compared with the other 2 testing sessions to confirm that background knowledge had increased over the course of the year. At each session students were also required to search a specialised database for the answers to a subset of the previous questions. (Only questions they did not know the answer to were used in the search task.) Wildemuth reports that over the first 2 sessions test scores and search performance both improve (but in principle improved search may have been due to increased familiarity with the database). Her analysis focused upon differences in the search strategies across the testing sessions.

We shall simplify (and at the same time extend) this method by using just one testing session and by explicitly relating the individual knowledge scores to the individual search performance. This enables us to firstly assess users' domain knowledge and then test the impact of this knowledge on search behaviour. More specifically, participants were given two sets of trivia questions each on a particular topic and asked to answer the questions unaided. Then, in the second phase of the procedure, they were given exactly the same questions again and asked to find the answers on the Web. The proportion of questions initially answered correctly on a topic thus indicates the level of domain knowledge and the knowledge test enables the identification of search questions which were and were not known by a participant. Because there are two different sets (topics) of questions it is possible to separate the specific effect of domain knowledge from any effect of general knowledge.

Our method also enabled study of search performance when participants already knew the answer. This acts as a comparison to searching when not knowing the answer and is of interest in its own right given the frequency that people search the Web to cross-check information [18]. Importantly, the design enables users to search the Web in an unrestricted manner, thus increasing the likelihood that the search behaviour observed is representative of performance outside the laboratory. The use of miscellaneous trivia questions as search tasks reduced the possibility that knowledge of a few specific Websites could determine search performance.[1]

Our main prediction then is that knowledge test score will correlate positively with search test score. This correlation should also be present when search performance is only measured on questions answered incorrectly in the knowledge test. The knowledge score for a topic should be more highly correlated with search performance on that topic than with search performance on the other topic.

Three ways in which domain knowledge might affect search

The primary purpose of this paper is to test whether domain knowledge predicts search performance on the Web and to introduce a new experimental design adapted from

---

[1] In fact 54% of the answers were found at Wikipedia but these pages were typically accessed direct from the search engine thus knowledge of the Website did not confer any great advantage.

Wildemuth [20]. However, a secondary objective is to provide some insights into *how* domain knowledge may shape search behaviour and to integrate this with existing theory. We anticipated that participants would rarely directly enter the URL of appropriate Websites, and this was confirmed. Therefore, within our experiment we suggest three ways in which knowledge of a topic could influence search.

Firstly, domain knowledge could affect the quality of the queries that are entered into the search engine. As discussed above, this is an issue that has been extensively investigated through the analysis of log transactions [9] and also by more qualitative studies of query formation within electronic databases [5,20].

Secondly, domain expertise could affect the links that are clicked on both from search engine summaries and when navigating more generally. Intuitively, it seems plausible that background knowledge would assist this process as greater knowledge would increase the ability of users to select links that are more likely to lead to the target information. Similarly, models of language comprehension (e.g. [11]) illustrate the importance of background knowledge when making inference judgements within text. Link following during navigation was discussed by [16] but has been most thoroughly addressed in studies of "information scent" (e.g. [3]). In this work the perceived relevance of a Web link to a user's goal is quantified. We suggest that this assessment of relevance will be facilitated by domain knowledge.

The third alternative is that domain expertise could affect the decision to leave a Webpage. That is, the decision to go to another page will not just be determined by the promise or "scent" of a link to another page but also by the likelihood that the answer can be found within the current page. Recent experimental work [17] has investigated the importance of such leaving decisions when allocating time across tasks, and shown that monitoring the gain of the current task is an important aspect of such decisions (as echoed in recent developments of models of online search [3].) Within this context the advantage of domain knowledge would be in the assessment of the likely gain from the current page. For example, an expert in pop music looking for information on the band "My Chemical Romance" might, appropriately, leave a Webpage about rap music more quickly than would a nonexpert.

To provide an indication of the type of search behaviour within our task we shall analyse search history data from a subset of our sample. Because it has been studied less than query formulation we shall focus more upon the way that experts and novices navigate around the Web than on specific details of search engine queries. Of course the actual queries entered might affect ongoing navigation behaviour and thus confound any effects of expertise (because domain experts may enter different queries into a search engine it is difficult to know if any differences in navigation behaviour are caused by the level of expertise or by the different set of results their queries produce). We try to be sensitive to this issue in discussing measures of navigation behaviour.

Each search task in our experiment required participants to retrieve a single fact from the Web. Because this information is relatively concrete and easy for readers to verify we do not anticipate that participants will often incorrectly leave Websites containing the target information. Thus, we predict that higher domain expertise will produce faster times to leave Webpages and, on average, each unsuccessful line of inquiry will be abandoned more quickly. The prediction is made stronger rather than confounded by the likelihood that experts will enter more effective queries into the search engine [5,7]. This is because if there were no differences in navigation performance we would expect the opposite result – better search queries should produce more promising search engine results meaning longer inquiries and a longer time before participants gave up each search. This argument enables us to be more confident that any negative correlation between domain expertise and time to appropriately give up is due to a difference in navigation performance rather than query formulation.

METHOD

Participants

Participants were 20 female and 14 male students from the authors' University with an average age of 21.7 years. Each was paid £6 for taking part. Detailed search behaviour data were collected from approximately a third of this group (7 females, 4 males). All participants reported searching the Web daily for their academic work and various leisure tasks. In the light of this no more refined measures of search expertise were taken.

Materials

Football (soccer) and pop music were selected as the two topics. These topics were deemed to be relatively independent domains of interest and a typical student population should provide a range of expertise on both subjects. Thus, we expected there to be some experts for each topic but across topics these experts would not necessarily be the same people.

Fifteen questions were constructed for each topic from a range of sources including experimenter knowledge and trivia Websites.[2] Each question required a single fact to be produced, typically a person's name or a band name or football club name. Sample questions are given in Table 1.

The difficulty level of the questions was set high (by experimenter judgement) to avoid ceiling effects and to provide sufficient examples of participants searching for

[2] None of the trivia Websites were searchable by google nor found by any participant during the experiment.

Music questions

- Which artist prevented My Chemical Romance from debuting at number 1 on the UK pop album chart?
- Which pop-rock group's original name was the Polar Bears?
- Who was Christina Aguilera's primary musical collaborator on her hit single "Hurt"?

Football questions

- Who was the only England player to score a hat-trick at international level in the 1970s?
- Which Irish defender moved on a free transfer to Wolves in 2002 after twelve years with Manchester United?
- Name either of the two players in the South African squad for the 2002 World Cup that were playing in the Premiership.

Table 1. Example trivia questions from each topic.

answers they didn't know. Similarly, questions were selected to prevent the search using the Web from being trivial. The criterion for this selection was that at the time of the experiment the correct answer did not appear within the first page of search results if the entire question was typed into Google. (Although the first page of search results sometimes provided a link to a page containing the correct answer.)

Design

All participants answered all questions in both the knowledge test and search test in a within-participants design. The proportion of questions correct for each topic on both the knowledge and search tests was recorded as was the time to find the answer during the search test. For the search test an item was only scored as correct if a Webpage containing the answer was found and saved.

Search behaviour was recorded for 11 participants. From these data, four measures were calculated for each question. "Number of pages" was the number of unique pages visited per question, including the initially returned list of summaries from a query. "Number of queries" referred to the number of unique queries entered into the search engine (Google) per question. The number of words in each unique query was also counted and the average number of words across queries for each item was the "Mean query length". We additionally computed "Number of paths abandoned" per question, i.e. the number of distinct paths (of any length) that were explored then abandoned. New paths were taken to start whenever a new query was entered or a second or subsequent Website from the search-returns page was selected. Clicking on further links was counted as lengthening the current path.

Procedure

After receiving instructions participants were presented with a practice question then the 30 trivia questions one at a time in a random order. The questions were presented at the top of a computer screen in a window that also contained a timer that counted down from 180 seconds. Participants were instructed to type the answer into a box below the question and then click on a button next to the box labelled "Enter". If they didn't know the answer they were told to guess or leave the box blank and click on Enter for the next question. The timer reset for each new question and if it reached zero participants were instructed to move on to the next question.

The knowledge test was completed when all 30 questions had been attempted. Participants were given further brief instructions and then completed the search test. After the practice question all of the questions were presented again in a different random order but otherwise following the procedure above. The key difference was that Internet Explorer was open in another window. See Figure 1 for a screenshot from the Search test.

Participants were requested to find a page on the Web that answered the question and then type the answer into the box as before within the 3 minutes. Even if they already knew the answer they still had to find a page that answered the question. After clicking Enter participants were then prompted to bookmark the last page and return to the homepage (Google.com) before receiving the next question. If they didn't find the answer participants had to keep searching for the entire 3 minutes return to the homepage and then move on to the next question. Apart from the

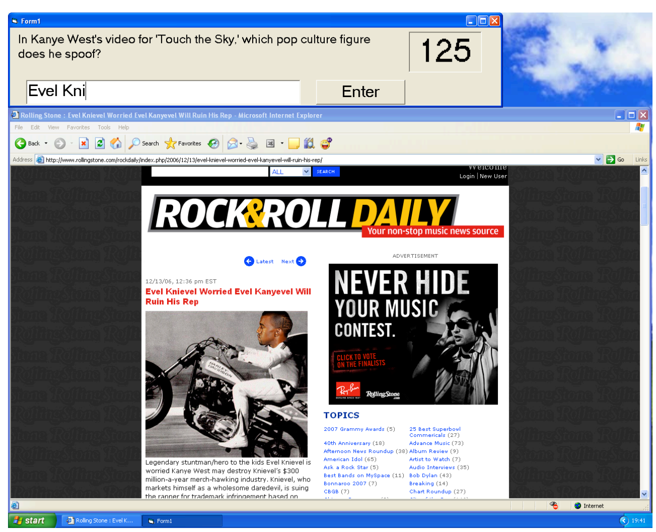

Figure 1. Screenshot from the search test. Questions were presented and answers entered using the window at the top. The larger window was used to search the Web.

constraint of starting from the homepage at the beginning of each question they were free to navigate the Web in any way they chose to find the answer.

RESULTS

The mean proportions of questions correctly answered on each topic were computed from the participant means for the knowledge test and for the search test and are presented in Table 2. As anticipated, for both topics a relatively small proportion could be answered by most participants from their knowledge alone. There was also no significant difference between the two topics in number of correct answers on the knowledge test, but it is worth noting that the variance in performance was considerably higher for the football questions than the music questions.

As predicted, during the search test a much higher proportion of questions were correctly answered. This effect was most pronounced for the music questions where the majority of participants scored highly and only 2 out of 34 participants correctly answered fewer than half of the questions. Performance on the football questions was also improved by the use of the Web but there were fewer correct answers than to the music questions, $t(33) = 5.96$, $p < .001$, $d = 1.02$, and a larger proportion did not score very highly (e.g. 15 of 34 failed to correctly answer half of the questions).

The mean proportion of correct answers during the search test to questions that participants answered incorrectly during the knowledge test is also given in Table 2. Again scores were higher for the music questions than for the football questions, $t(33) = 3.87$, $p < .001$, $d = .66$. This

	Football		Music	
	M	*SD*	*M*	*SD*
Knowledge score	.17	.22	.18	.13
Search score (all items)	.54	.17	.68	.16
Search score (not-known items)	.51	.16	.62	.16
Search time (all items)	122.42	23.22	104.45	23.64
Search time (not-known items)	129.07	16.70	116.34	23.71
Search time (items correct)	75.18	15.50	69.86	18.08

Table 2. Mean proportion correct and mean times on knowledge and search tests.

Note: All times given in seconds. Knowledge score and search score are the proportion of items correctly answered. Not-known items are those answered incorrectly on the knowledge test.

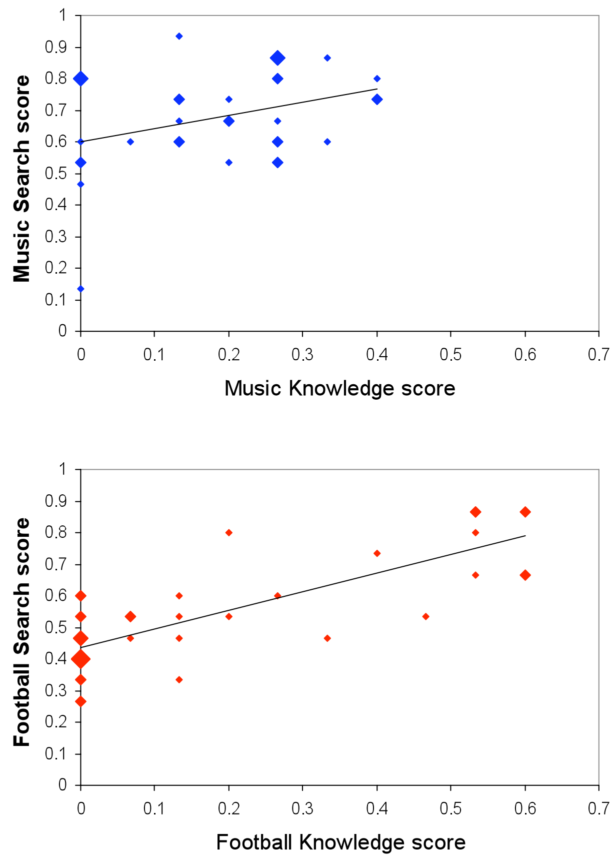

Figure 2. Correlations between knowledge and search performance for football and music scores on all items.

Note: The size of the data point indicates the number of participants represented. The number of participants at each data point ranged between 1 and 4 for the football questions and 1 and 3 for the music questions. Larger data points represent more participants.

measure enables an assessment of the influence of knowledge of the topic on search performance to be made without being confounded by the specific knowledge of the answer. However, if we assume those questions that are more easily answered without the Web are also more easily answered with the Web then it should be noted that this measure will lead to a conservative estimate of the influence of expertise. That is, experts' search performance will only be assessed on the more difficult questions as they will have already correctly answered the easier questions.

To investigate the relationship between performance on the knowledge test and the search test Pearson's r correlations were computed on the individual participant means. In Figure 2 the proportion correct on the knowledge test is plotted against the proportion correct on the search test for each topic. Eyeballing the data indicates a positive correlation for both topics with a greater range of scores for the football questions than for the music questions. This was confirmed by the analysis as knowledge scores correlated with the subsequent search scores for music questions, $r(34) = .35$, $p < .05$, and, more strongly, for football questions, $r(34) = .76$, $p < .001$.

There was also a significant correlation between the knowledge scores on the two different topics, $r(34) = .48$, $p < .01$. This indicated that those people who knew more about football were more likely to also know more about music. Thus, we might expect that knowledge on one topic might correlate to some extent with search performance on the other topic. This was the case as football knowledge correlated with music search score, $r(34) = .42$, $p < .05$, and music knowledge correlated with football search score, $r(34) = .41$, $p < .05$. Therefore, it is not possible to completely discount the hypothesis that the correlations in Figure 2 are due to some general factor rather than expertise in the topic domain *per se*. For example, a better memory might aid retention of general knowledge and boost searching performance. Alternatively, spending more time online might improve general knowledge and enhance searching skill.

To address these concerns two regression analyses were conducted one with football search score as the dependent variable and the other using music search score. Independent variables were the knowledge scores on the two topics. Both regression equations were significant, football: $R^2 = .583$, $F(2, 31) = 21.64$, $p < .001$; music: $R^2 = .20$, $F(2, 31) = 3.98$, $p < .05$, and the results are given in Table 3. The findings indicate that football knowledge score was a significant predictor of football search score but music knowledge predicted no additional variance in football search, arguing against an effect of general knowledge (or at least arguing against such an effect being comparable in size to the effect of football knowledge). For

	B	*SE*	*β*	*p*
Football search score (all items)				
Football knowledge	.57	.10	.73	<.001
Music knowledge	.08	.17	.07	.63
Music search score (all items)				
Football knowledge	.24	.13	.33	.08
Music knowledge	.23	.22	.19	.30
Football search score (not-known items)				
Football knowledge	.35	.13	.49	.01
Music knowledge	.05	.21	.04	.83

Table 3. Regression analyses on search scores

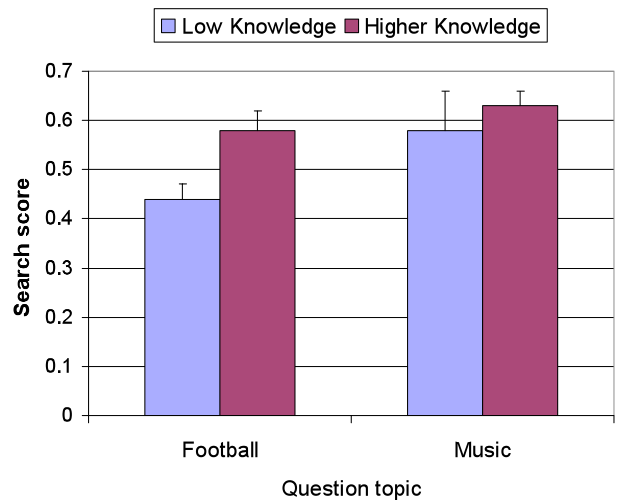

Figure 3. Higher and low knowledge participants mean search scores for not known questions

music search score the findings were more equivocal with neither variable reliably predicting search score.

Search performance for not-known items

Search performance was also investigated using questions answered incorrectly on the knowledge test. This provided a more stringent test of the role of expertise in assisting search as participants did not know the specific answer to these questions before carrying out the search. The analysis found that for the football questions knowledge score correlated with search score, $r(34) = .51$, $p < .01$. No other correlations were significant. The failure to find a correlation between music knowledge and football search performance ($r = .27$) supports the suggestion that it was knowledge of football that improved the ability to find information on football rather than some other general factor. But, for confirmation, this was investigated using the same regression analyses as above. The regression equation for football search score was significant, $R^2 = .26$, $F(2, 31) = 5.38$, $p < .05$, and the results are included in Table 3. Once again domain knowledge predicted football search performance but music knowledge did not. The regression equation for music search score was not significant and thus the analyses are not included in Table 3. We shall return to the absence of a correlation between music knowledge and music search score in the discussion.

Showing that domain knowledge assisted search performance even when participants didn't know the answer is a principal goal of this paper. Therefore, we provide a more graphic display of this effect by comparing participants with very little knowledge to those who had some knowledge. For each topic area participants were split into two groups, the low knowledge group scored 0 on the knowledge test (Football, $N = 15$; Music, $N = 8$) and the higher knowledge group (Football, $N = 19$; Music, $N = 26$)

	Football				Music			
	All items		Not-known items		All items		Not-known items	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Pages visited	4.28	.76	4.47	.50	4.39	.82	4.75	.82
Search queries	1.87	.43	1.96	.39	2.05	.41	2.16	.39
Mean query length	4.46	1.28	4.58	1.13	3.92	.92	4.13	1.00
Paths abandoned	2.76	.66	2.92	.51	2.92	.62	3.19	.79
Time per page	24.99	3.94	26.28	3.04	22.37	3.26	23.77	3.54
Time per query	61.24	15.39	63.88	17.58	52.77	14.23	54.32	15.38
Time to abandon path	45.02	9.27	45.92	8.53	40.12	8.46	39.87	9.13

Table 4. Measures of search queries and pages visited.

Note: All times given in seconds.

answered at least one question correctly.[3] Figure 3 shows that during the search test the higher knowledge group found a higher proportion of the items they didn't previously know the answer to than the low knowledge group. This effect was significant for the football questions, $t(32) = 2.88$, $p < .05$, $d = .72$, but not for the music questions, $t(32) = .69$, $d = .17$.

Search times

As reported in Table 2 a considerable proportion of the search items were not found within the 3 minute time limit. In light of this, two different methods were used to analyse search times. The first method was to assign a time of 180 seconds for any item answered incorrectly either through providing the wrong answer or not answering within the 3 minute time limit. This method meant the timing data was to some extent derivative of the proportion answered correctly which were analysed above. This was borne out by the direction of effects and pattern of significance which was the same as reported above for every comparison and correlation apart from the correlation between music knowledge and music search time for all items which was not significant ($r = -.26$).

For the second method the participant means were calculated using only the items correctly answered during the search test. Analyses of these data found no difference in search times for the two topics and only one reliable correlation – between football knowledge and football search time, $r(34) = -.37$, $p < .05$. This correlation was negative, indicating that participants who knew more about football found the answers on the Web more quickly.

[3] A median split using knowledge score produced an identical pattern of significance but was inappropriate for statistical reasons, see [14])

Measures of search behaviour

Descriptive statistics for the searches completed by the 11 participants whose search behaviour was logged are given in Table 4. Means and standard deviations were calculated from the individual participants' mean scores for each measure. Where participants made a typing error and subsequently corrected for it in an otherwise identical search query (8% of queries) the erroneous query was excluded. Search time was divided by number of pages visited to provide an index of time spent on each page. This corrected for the between-item variance in time spent searching. Time per search query was similarly calculated. Because participants did not give up any searches on a sizeable number of items the time to abandon a path was computed using Time/(Paths abandoned + 1).

When all items were included, mean length of query was longer for football than for music, time per page was longer for football than for music and time per query was longer for football than for music, $2.38 \leq t\text{s}(10) \leq 2.83$, $p\text{s} < .05$, $.49 \leq d\text{s} \leq .73$. These differences were not reliable when only not-known items were considered. Time to abandon a path was longer for football than for music for all items and for not-known items, $2.28 \leq t\text{s}(10) \leq 2.40$, $p\text{s} < .05$, $.55 \leq d\text{s} \leq .69$. No other differences between question topics were significant.

Search Navigation

Correlations were computed between knowledge scores for both topics and all seven measures of search behaviour shown in Table 4 for each topic. Both when all items, and when only not-known items were analysed the only significant correlation was between football knowledge score and time per page for all the football items, $r(11) = -.70$, $p < .05$. This relationship is shown in Figure 4. Regression analysis was not appropriate due to the small sample size [19]. Thus, participants with high football knowledge spent less time on each page visited when searching for an answer to the football questions. That the

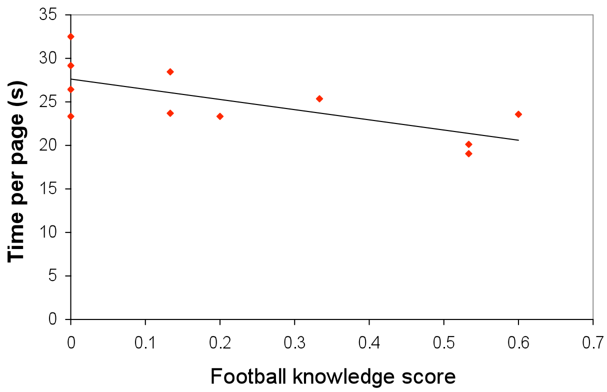

Figure 4. Correlation between football knowledge and time per page when searching for football items.

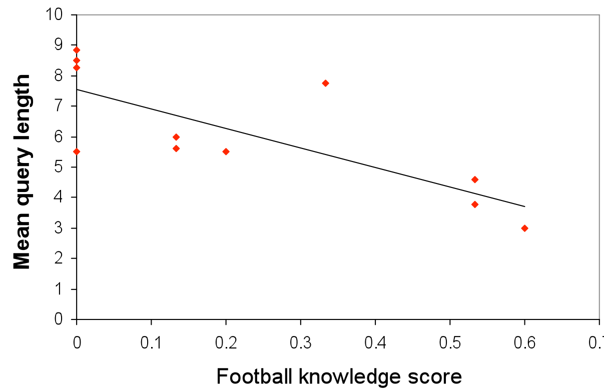

Figure 5. Correlation between football knowledge and mean query length when searching for answers to two difficult football questions.

effect was significant with just 11 participants is remarkable and reflects the unusually large effect size [2].

With such a small sample size it may be worth considering notable effects alongside significant effects. In addition to the significant negative correlation with time per page, there was a large correlation between football knowledge and time to abandon a path for all football items, $r(11) = -.44$. This indicated that for the football items experts abandoned searches more quickly. There were also negative correlations between football knowledge and number of pages, number of queries and paths abandoned, $-.52 \leq r\text{s}(11) \leq -.47$. But these results were unsurprising given the correlations between time spent searching and domain expertise reported above.

All these effects were specific to the football domain and were not replicated for the music questions. However, this was predictable given the overall failure to find a clear relationship between music knowledge and music search performance. There were large correlations between football knowledge and time to abandon a path for all music items, $r(11) = -.49$, and for items incorrect in the music knowledge test, $r(11) = -.59$. Music knowledge did not correlate with any search behaviour for the football questions.

Query Formulation

Within the football search behaviour there was also a marginally significant correlation between mean query length and football knowledge when all items were included, $r(11) = -.57$, $p = .07$ and when not-known items were analysed this correlation was still sizeable, $r(11) = -.46$. To control for the amount of time spent searching we selected the two most difficult items to find for both music and football (each of the four items were failed by >85% participants). This meant most participants would have spent an equal time of 3 minutes searching for each item. All correlations described above were then computed on this new dataset excluding the redundant variables time per page and time per query. There was a significant correlation between football knowledge and mean query length for the two football items, $r(11) = -.77$, $p < .01$ and no other correlations were significant.

This result is illustrated in Figure 5 and means that participants with a high knowledge of football tended to use shorter queries than participants with a low knowledge of football. This effect appeared to be domain specific as high knowledge in one area (either football or music) did not correlate with shorter queries when searching in the other area.

DISCUSSION

Perhaps the key result from this experiment was the finding that football knowledge predicted search performance on football questions that participants did not already know the answer to. Thus, our novel methodology has successfully demonstrated that domain knowledge can help users find information on the Web.

When including questions that participants already knew the answer to within the analysis of search performance, topic knowledge correlated with search performance for both music and football questions. This result is less theoretically interesting as participants were able to use the actual answers to aid their search. Nonetheless, this is not without applied relevance as the Web is frequently used to cross-check or verify information [18].

The results are in general much weaker for the music questions, but there is a ready explanation for this in the much lower variance of scores for the music knowledge test compared to the football knowledge test. This effect was compounded by the restricted range of scores on the search test for the music questions where almost all (32 out of 34) participants correctly answered more than half of the music questions. The problem is perhaps shown most clearly in Figure 2 where the scores are spread across both axes for the football questions and clustered around the centre for the music questions. Therefore, irrespective of the relationship between knowledge and search performance, the music questions did not discriminate well between different levels of knowledge or between different levels of search performance. This meant the power of any test of the relationship between knowledge and search was severely restricted. It is possible that the structure of information on the Web means that for some topics it is harder to develop answerable questions that are not easy to find. However, this issue is beyond the scope of this article and perhaps the most likely explanation for these problems is simply that the music questions were not well chosen.

The data collected on search behaviour also yielded some interesting trends. The most striking of these was the large negative correlation between football knowledge score and the mean time spent on each page when searching for football items. This effect indicated that participants more knowledgeable in football spent less time on each page when searching for information about football. Like the correlation between knowledge and overall search performance this was domain specific and did not occur for the music questions. There was a similar but nonsignificant relationship between knowledge and time to abandon a path for the football items in that football experts spent less time on each individual line of inquiry when searching. That is, more knowledgeable participants returned to a previous set of search engine results or entered a new query more quickly than less knowledgeable participants, indicating they took less time to conclude that a particular search was likely to be unsuccessful. As noted earlier it is difficult to attribute this trend to improved query formulation with expertise because better queries would be expected to produce more promising lines of inquiry that would therefore take more time to abandon rather than less. Instead, these correlations appear to reflect differences in navigation that result from background knowledge.

The trend for experts to give up each search more quickly supports the suggestion that domain knowledge helps participants to evaluate a Webpage faster and thus make a quicker decision to abandon unpromising Websites. Time spent on each page presumably also reflects this judgment, but the time per page measure also includes pages which the participant did not abandon, but rather navigated forwards from (to the extent that this distinction can be made – our statistic of abandoned paths assumes it can be). This is interesting because the correlation between time per page and domain expertise was very large and significant, perhaps because it represented the sum of two navigational effects. More knowledgeable participants may have been able to abandon a line of inquiry more quickly and also more able to quickly select and follow a link from a Webpage.

This explanation is aligned with Fu and Pirolli's model of Website navigation [3]. Unlike this article their focus is largely upon producing quantitative fits to navigation data. However, the qualitative patterns described here are consistent with the SNIF-ACT 2.0 model of Website navigation. The model sequentially evaluated links on a Webpage and used this assessment to guide decisions to either leave a Webpage or follow a link. The evaluation process involved the formation of associations between the link and the search goal and should therefore be assisted by background knowledge [11]. If domain knowledge enhances the process of link evaluation the model would predict that both link following and giving up behaviour would be affected. As a secondary point it is worth noting that the methodology deployed here provides a readymade index of background knowledge that could be used for modelling the effects of domain knowledge on search.

As with the general correlations between knowledge and search scores these results were domain specific and were not apparent for the music questions. However, the trend towards a correlation between football knowledge and time to give up a line of inquiry when searching for music items indicated that general background knowledge could also predict the tendency to give up a search.

The other notable finding from the search behaviour was that football knowledge correlated with mean query length when searching for answers to the football questions. Again there was no equivalent correlation for music questions and again this could be a consequence of the problems associated with the music questions discussed above. The result for the football questions indicated that participants with higher knowledge on average entered shorter queries into the search engine. This effect is true across all questions, and also when success on the questions is controlled for. When participants already knew the answer to a question this result might be simply explained as high knowledge participants using the one or two word answer as a search term. However, the data indicate that the finding extended to questions that participants answered wrongly during the knowledge test. Given that participants more knowledgeable in the topic domain would be expected to know more potential search terms this result is particularly striking. It's hard to think of a cruder index of the quality of a query than simple number of words, and yet this crude index shows a clear quantitative effect of background knowledge on query formulation. This effect supports the finding within the Library and Information Sciences literature that query formulation will be affected by domain knowledge. The use of short queries is also consistent with the general tendency to use shorter queries when searching on the Web relative to traditional information retrieval systems [9].

A strength of the work presented here is that few constraints were placed upon participants when they interacted with the Web increasing the generalisability of our findings. Against this, it must be admitted that our time constraint of 3-minutes per question was an experimental convenience to enable measurable effects in limited time rather than a realistic estimate of how long searchers are prepared to search for valuable information. Would less restricted times exaggerate or diminish the difference between more and less successful searchers? One might speculate that with unlimited time everyone would find every answer, but we suspect that there will be a strong influence of cognitive factors on persistence. Indeed, we think that this is an important topic for future theoretical and experimental research.

Whatever the successes and limitations of our current experiment, its balance of laboratory constraint and free search behaviour, we are sure that some more tightly controlled experimentation will also be required to confirm the suggestions presented here about *how* background knowledge influences Web search performance.

To conclude, within this article we have introduced an adapted design that provides a new and, we like to think, elegant method for exploring the role of domain expertise when searching on the Web. Results produced using this method demonstrated that even when participants did not know the specific answer, background knowledge in the topic assisted search performance. Exploratory analyses suggested that background knowledge improved both search engine query formulation and general navigation of the Web.

We began this article by wondering whether the ready availability of information threatened the societal value of expertise. It has become a commonplace in the popular media to assume that when technology provides easy answers, the associated intellectual skills decline (e.g. [4,8]). Is there any advantage in remembering facts when they can be searched for? Our experiment suggests there is – easier access to the rest of the field of knowledge. The finding that knowing more helps you find more raises the intriguing possibility that in a certain sense the technologies of Web search might amplify rather than diminish the effects of traditional 'in the head' expertise.

REFERENCES

1. Bhavnani, S.K. (2001). Important Cognitive Components of Domain-Specific Search Knowledge. *Proceedings of TREC'2001*, 571-578.
2. Cohen, J. (1988). *Statistical power for the behavioral sciences* (2nd Edition). Hillsdale, NJ: Erlbaum.
3. Fu, W.-T., Pirolli, P. (in press), SNIF-ACT: A Model of Information-Seeking Behavior in the World Wide Web. *Human-Computer Interaction*.
4. Greenfield, S. (2003). *How tomorrow's technology is changing the way we think and feel.* London, Penguin.
5. Hembrooke, H.A., Gay, G.K., Granka, L.A. (2005), "The effects of expertise and feedback on search term selection and subsequent learning", *Journal of the American Society for Information Science and Technology, 56(8)*, 861-71.
6. Hölscher, C., & G. Strube. (2000). Web search behavior of internet experts and newbies. *Computer Networks, 33(1)*, 337–346.
7. Hsieh-Yee, I. (1993). Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers. *Journal of the American Society for Information Science and Technology, 44(3)*, 161–174.
8. James, C. Helplessly advanced. On clever machines. http://www.clivejames.com/point-of-view/helplessly-advanced.
9. Jansen, B.J., & Pooch, U. (2000). A review of web searching studies and a framework for future research. *Journal of the American Society of Information Science and Technology, 52(3)*, 235–246.
10. Jansen, B.J., & Spink, A. (2003). An analysis of Web documents retrieved and viewed. *Proceedings of the 4th International Conference on Internet Computing.* 65-69, Las Vegas, NV.
11. Kintsch, W. (1998). *Comprehension A paradigm for cognition.* New York: Cambridge University Press.
12. Kuhlthau, Carol C. (1991). Inside the Search Process: Information Seeking from the User's Perspective. *Journal of the American Society of Information Science, 42*, 361-371.
13. Lazonder, A., Bieiemans, H., & Wopereis, I. (2002). Differences between novice and experienced users in searching for information on the World Wide Web. *Journal of the American Society for Information Science, 51*, 576–581.
14. MacCallum, R.C., Zhang, S., Preacher, K.J., & Rucker, D.D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7,* 19-40.
15. Marchionini, G. (1997). Information seeking in electronic environments. Cambridge, Cambridge University Press.
16. Mat-Hassan, M., & Levene, M. (2005). Associating search and navigation behavior through log analysis. *Journal of the American Society for Information Science and Technology, 56(9)*, 913–934.
17. Payne, S. J., Duggan, G. B., & Neth, H. (2007). Discretionary task interleaving: Heuristics for time allocation in cognitive foraging. *Journal of Experimental Psychology: General, 136(3)*, 370-388.
18. Sellen, A.J., Murphy, R., & Shaw, K.L. (2002). How knowledge workers use the Web. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 227–234, Minneapolis, MN.
19. Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistics (2nd Edition)*. Cambridge, MA: Harper & Row.
20. Wildemuth, B. M. (2003). The effects of domain knowledge on search tactic formulation. *Journal of the American Society for Information Science, 55*, 246-258.